Pacific Association Player Report

How the Sonoma Stompers dominated the league at the plate and on the mound

Tim Livingston - OPS 804

2016-10-29

Introduction

This report is to determine the range of hitter performances accross the Pacific Association based on an advanced metric that have been developed to account for the most accurate measure of how well a player is doing relative to the league. We will use one metric for a hitter and one for a pitcher. For hitters, we are going with Weighted Runs Created Plus (wRC+). For pitchers, Fielding Independent Pitching (FIP). With the Sonoma Stompers winning the championship this season, I wanted to take a look at how well our team did relative to the rest of the teams in the league, and to see whether or not this is a performance that can be repeated or not.

wRC+ is based on a percentage scale, where 100 is considered league average. Any points above mean that players are that much better and vice versa for players lower. For example, if a player has a wRC+ of 130, that means he's 30% better than the league average player. If his WRC+ is 75 instead, he's 25% worse than the league average player.

FIP is a lot like Earned Run Average (ERA), but only accounts for what a pitcher can control: strikeouts, walks, and home runs. This is to make sure we take any bad luck out of the equation thanks to poor defense, putting the pitchers on a more level playing field.

To go through this data, I did a couple of arbitrary cutoffs to account for the best possible data. First, for wRC+, I set a minimum of 100 plate appearances (PA), close to the median of 101. 100 is more of a round number and can be used going forward as a good measure to gauge players who were a part of the league for a good amount of time. This year, the league average wRC+ was 88, with a median of 92. The data for the hitting report is as follows:

Batter	Team	wRC
Nick Oddo	PIT	131
Jordan Hinshaw	PIT	118
Mike Taylor	PIT	82
David Kiriokos	PIT	104
Rich Mejia	PIT	78
Thomas Shull	PIT	96
Andrew Rubalcava	PIT	105
John Contreras	PIT	121
Jordan Yallen	PIT	34
Brandon Williams	PIT	82
Tyler Nordgren	PIT	142
Sammy Ayala	PIT	116
Ian Hagenmiller	PIT	54
Derrick Fox	SON	105
Daniel Baptista	SON	112
Joel Carranza	SON	141
Mark Hurley	SON	113
Matt Hibbert	SON	110

Batter	Team	wRC
Mason Morioka	SON	123
Yuki Yasuda	SON	121
Caleb Bryson	SON	129
Eddie Mora-Loera	SON	102
Chaz Meadows	SON	80
Randy Santiesteban	SON	147
Ethan Szabo	SON	121
Zack Pace	SR	119
Chase Tucker	SR	120
Jake Taylor	SR	132
Maikel Jova	SR	74
Danny Gonzalez	SR	81
Brent Gillespie	SR	109
Ricky Gingras	SR	66
Johnny Bekakis	SR	120
Kyle Adie	SR	159
Darian Sandford	VAL	88
Tim Williams	VAL	123
Lydell Moseby	VAL	114
Gerald Bautista	VAL	114
Gadiel Baez	VAL	78
Joseph Hicks	VAL	86
Wilfredo Petit	VAL	80
P.J. Phillips	VAL	60
Alian Silva	VAL	58
Aaron Brill	VAL	60
Jackson Valera	VAL	164
Pedro Perez	VAL	85
Cyle Figueroa	VAL	75

And the pitchers, who had a league-wide mean FIP of 4.46 and a median of 4.01, giving us 51 pitchers that were at or above the median value for innings pitched, which was just above 14:

Pitcher	Team	FIP
Patrick Conroy	SR	3.095
Celson Polanco	SR	3.524
David Dinelli	VAL	4.157
Roman Gomez	VAL	3.833
Taylor Thurber	SON	2.345
Nick DeBarr	SR	4.224
Chris Cummins	PIT	2.771
Marquis Hutchinson	VAL	4.932
Dylan Brammer	PIT	3.509
Patrick Barnett	SR	2.672
Gregory Paulino	SON	4.445
Mike Jackson Jr.	SON	3.597
Corey MacDonald	PIT	4.225
Kida De La Cruz	VAL	4.550
Juan Espinosa	SON	3.309
Sean Conroy	SON	4.715
Oliver Garcia	SON	3.713

Pitcher	Team	FIP
Michael Ormseth	SR	2.534
Justin Lawrence	PIT	3.707
Julian Esquibel	SR	4.274
Rob Ellis	PIT	2.267
Jose Flores	SON	3.696
Wendell Floranus	VAL	4.233
Jeff Lyons	PIT	3.839
Josh Evans	VAL	2.158
Joe Watson	VAL	4.875
Brent Adheen	VAL	4.702
Garrett Granitz	PIT	2.929
Marquis Pettis	PIT	4.622
Brandon Warner	SR	3.873
Hector Ortiz	VAL	6.467
J R Bunda	SR	2.004
Erik Gonsalves	SON	3.619
Clint Manzo	SR	4.012
Michael Kershner	SR	4.953
Ricky Schafer	PIT	5.143
Joe Mello	PIT	2.623
Carlos Pinales	VAL	2.730
Martin Cronin	SON	4.006
Matt Picucci	SON	3.568
Justin Hertzmann	VAL	3.043
Joe Lewis	PIT	3.643
Eric Mozeika	SON	3.150
Nate Gercken	SR	6.067
Miles Richard	PIT	5.479
Austin Delmotte	SON	3.285

As a general rule, players who played with multiple teams will be considered with the team they played the most for this year using the given parameters, as the data they gathered in the small sample with their second team wouldn't affect the overall performance.

I ran descriptive stats for the Hitters:

count	sum	min	max	mean	median	range	q1	q3	iqr	sd	var	kurt	skew
47	4832	34	164	103	109	130	80	120	40.5	29	826	-0.531	-0.1315

Now the descriptive stats for the pitchers:

count	sum	min	max	mean	median	range	q1	q3	iqr	sd	var	kurt	skew
46	175.117	2.004	6.467	3.81	3.71	4.463	3.11	4.4	1.2935	1	1	-0.1029	0.3896

Next, the ANOVA analysis, using numerous tests to determine what the data set looks like on a team by team basis. The given hypothesis is that hitting performance was similar amongst all teams in the league. Same with pitching. First, the hitters (NOTE: Sorry this isn't in a nice table like the previous data. Will have to work this bug out in future reports.):

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Team         3   4931  1643.7   2.136   0.11
## Residuals   43  33084   769.4
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = fit)
##
## $Team
##                diff        lwr       upr     p adj
## SON-PIT  19.846154  -9.828752 49.521060 0.2933036
## SR-PIT   11.735043 -20.408967 43.879053 0.7639363
## VAL-PIT  -6.000000 -35.075351 23.075351 0.9456463
## SR-SON   -8.111111 -40.798439 24.576216 0.9103084
## VAL-SON -25.846154 -55.521060  3.828752 0.1075631
## VAL-SR  -17.735043 -49.879053 14.408967 0.4614100
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  3  0.7279  0.541
##       43
```

To recap:

-ANOVA shows that we can not reject the null hypothesis via the given values for hitters. This indicates that perhaps the teams were more similar than first believed across the board. Even at a 90% confidence interval, this would not have mattered.

-Tukey analysis showed Pittsburg and Vallejo as the closest in hitting performance, with San Rafael and Sonoma second. Neither hit the 95% confidence interval, however.

-Levene test showed a similar finding, although the homogeneity of the data is very staggered, as you'll see once we get to the graphs.

Now the pitchers, with a similar hypothesis test:

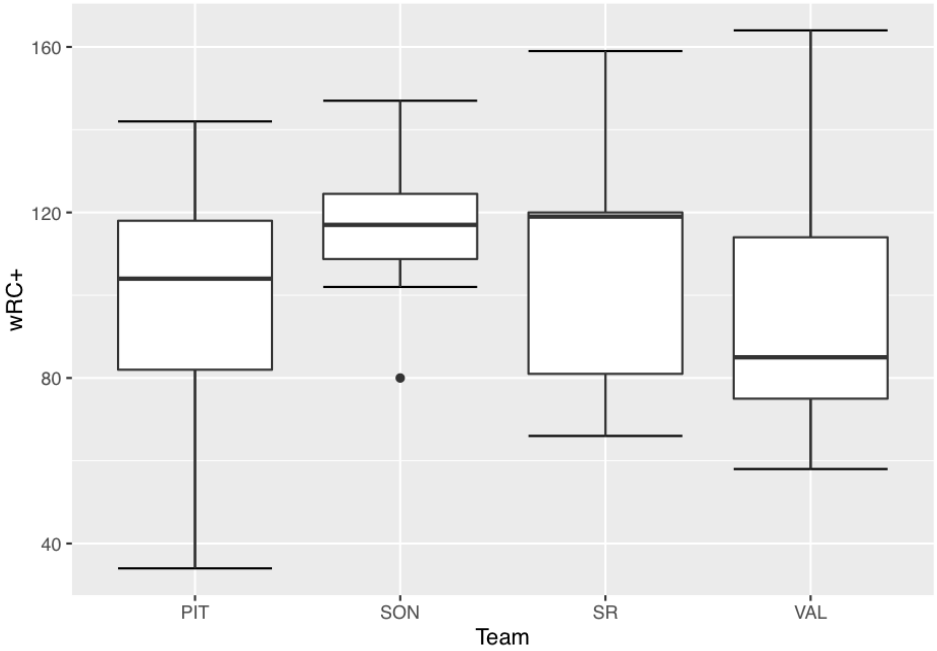
```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Team         3   1.84  0.6136   0.599  0.619
## Residuals   42  43.05  1.0250
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = fit)
##
## $Team
##                 diff        lwr       upr     p adj
## SON-PIT -0.10908333 -1.2146912 0.9965246 0.9934649
## SR-PIT   0.01861364 -1.1118425 1.1490698 0.9999688
## VAL-PIT  0.42297727 -0.7074788 1.5534334 0.7497071
## SR-SON   0.12769697 -1.0027591 1.2581531 0.9902766
## VAL-SON  0.53206061 -0.5983955 1.6625167 0.5935897
## VAL-SR   0.40436364 -0.7504061 1.5591334 0.7853736
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  3  1.3988 0.2565
##       42
```

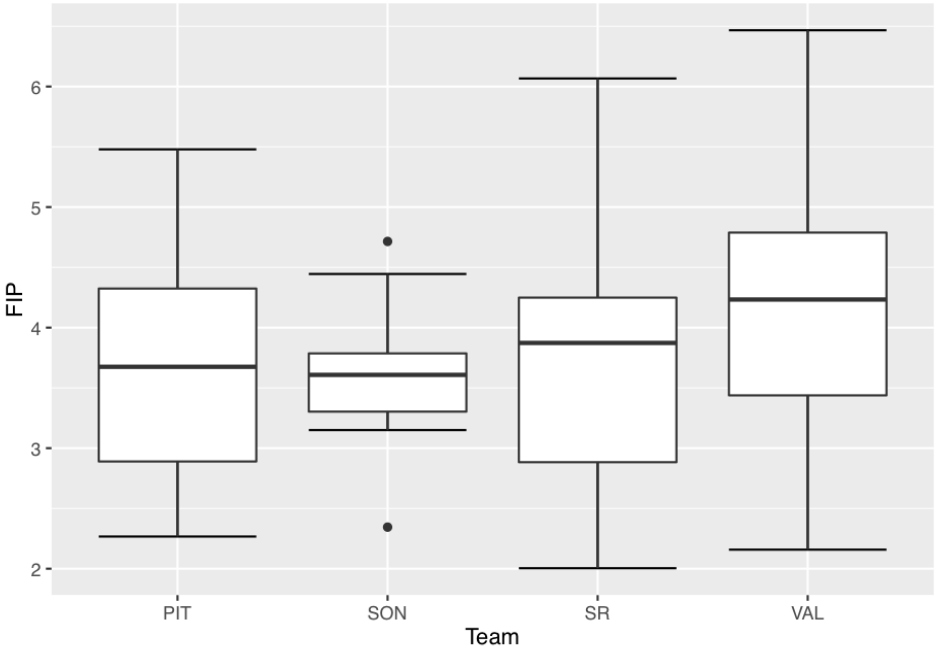
Again, similar results, but even more so for pitching, as the high offense environment led to incredible variance amongst our pitchers.

Finally, box and whisker plots were created for both wRC+ and FIP on a team by team basis. This is to determine how much variance there was between players on a given team and to also determine if a team performed better on the whole than another. For the hitters, this is what you see:



As you can see, the Stompers performed extremely well, with a tight grouping of data relatively higher than the mean of the data, which was about 103. Most teams did have good performers, but at the same time, didn't have as many that performed as well as Sonoma did. This is an extremely good indicator of our team's performance, and is a major data point to show just how well the team performed as a whole.

Now for the pitchers:



Again, a very tight grouping that allows us to feel good about our overall performance, as our IQR tops out at about the mean FIP of the league and goes down from there. Another fantastic piece of evidence about how well our team did on the mound.

The last two graphs really tell the story about how well we did as a team. Overall, having the type of pitching we did along with the hitting we did led us down the championship path. We should be very happy with our performance this season, and the numbers back it up.